

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-81746

(43) 公開日 平成9年(1997)3月28日

(51) Int.Cl.⁶

G 0 6 T 7/00
17/00

識別記号

庁内整理番号

F I

G 0 6 F 15/62

技術表示箇所

4 1 5
3 5 0 A

審査請求 未請求 請求項の数38 F D (全 23 頁)

(21) 出願番号 特願平7-256772

(22) 出願日 平成7年(1995)9月8日

(71) 出願人 000001889

三洋電機株式会社

大阪府守口市京阪本通2丁目5番5号

(72) 発明者 松本 幸則

大阪府守口市京阪本通2丁目5番5号 三

洋電機株式会社内

(72) 発明者 寺崎 肇

大阪府守口市京阪本通2丁目5番5号 三

洋電機株式会社内

(72) 発明者 杉本 和英

大阪府守口市京阪本通2丁目5番5号 三

洋電機株式会社内

(74) 代理人 弁理士 吉田 研二 (外2名)

最終頁に続く

(54) 【発明の名称】 二次元表示画像生成方法

(57) 【要約】

【課題】 視点変更映像、画像領域の自然な拡大・縮小、所期の画像領域の切り出しを精度よく行うことが困難だった。

【解決手段】 入力映像から、被写体の画面平面上の動きを抽出する(工程1)。三次元空間における被写体の現実の動きを算出する(工程2)。入力映像はもとの三次元的な動きの投影であるから、複数の代表点の動きから三次元表示の動きが逆変換によって求められる。この結果、被写体の座標が判明し、被写体の奥行き情報が獲得される(工程3)。この三次元情報により視点変更映像等が生成される(工程4)。

工程1: 二次元動き情報の抽出

工程2: 三次元動き情報の算出

工程3: 奥行き情報の獲得

工程4: 画像の生成

3/13

【特許請求の範囲】

【請求項1】 処理の対象となる元の映像から該映像の奥行き情報を抽出し、この情報に従って二次元表示用の画像を生成することを特徴とする二次元表示画像生成方法。

【請求項2】 請求項1に記載の二次元表示画像生成方法において、複数の映像フレーム間における映像各部位の二次元的な位置の変位を検出し、この情報から前記奥行き情報を抽出することを特徴とする二次元表示画像生成方法。

【請求項3】 請求項2に記載の二次元表示画像生成方法において、前記複数の映像フレームは異なる時刻に撮影されたものであり、前記二次元的な位置の変位量に基づいてこれら複数の映像フレームを選択することを特徴とする二次元表示画像生成方法。

【請求項4】 請求項3に記載の二次元表示画像生成方法において、前記映像各部位の二次元的な位置の変位量を統計的に処理し、その処理結果に基づいて前記複数の映像フレームを選択することを特徴とする二次元表示画像生成方法。

【請求項5】 請求項4に記載の二次元表示画像生成方法において、前記変位量の統計的な処理は、前記映像各部位の動きベクトルの分散の導出であり、この分散が所定値よりも大きくなるように前記複数の映像フレームを選択することを特徴とする二次元表示画像生成方法。

【請求項6】 請求項5に記載の二次元表示画像生成方法において、前記分散が所定値よりも大きくなるような複数の映像フレームを選択することができない場合には、二次元表示用画像の生成を中止することを特徴とする二次元表示画像生成方法。

【請求項7】 請求項2～6のいずれかに記載の二次元表示画像生成方法において、前記二次元的な位置の変位から前記映像各部位が現実の三次元空間において占める相対位置関係を導出し、その結果に従って前記奥行きを決定することを特徴とする二次元表示画像生成方法。

【請求項8】 請求項2～6のいずれかに記載の二次元表示画像生成方法において、前記二次元的な位置の変位から前記映像各部位の三次元的な動きを算出し、この動きから三角測量の原理によって前記映像各部位の三次元空間における位置座標を算出し、その結果に従って前記奥行きを決定することを特徴とする二次元表示画像生成方法。

【請求項9】 請求項2～8のいずれかに記載の二次元表示画像生成方法において、

基準となる映像フレームに代表点を設定し、他の映像フレームにおける前記代表点の対応点を求め、これら代表点と対応点の位置関係を求めることによって、前記映像各部位の二次元的な位置の変位を認識することを特徴とする二次元表示画像生成方法。

【請求項10】 請求項9に記載の二次元表示画像生成方法において、前記基準となる映像フレームと前記他の映像フレームが異なる時刻に撮影されたものであるとき、前記代表点と前記対応点の位置関係から、さらに他の映像フレームの対応点の位置を予想し、対応点を探索すべき領域を限定することを特徴とする二次元表示画像生成方法。

【請求項11】 請求項9、10のいずれかに記載の二次元表示画像生成方法において、前記代表点のうち、幾何的に特徴のある映像領域に関連するものについては、それら代表点の対応点の関連する映像領域についても前記幾何的な特徴が維持されるよう、前記対応点の位置を調整することを特徴とする二次元表示画像生成方法。

【請求項12】 請求項11に記載の二次元表示画像生成方法において、前記幾何的に特徴のある映像領域は、直線を含む領域であることを特徴とする二次元表示画像生成方法。

【請求項13】 請求項9に記載の二次元表示画像生成方法において、前記他の映像フレームにおける特定点近傍の映像領域と、前記基準となる映像フレームの代表点近傍の映像領域間で映像の類似性を評価し、この評価結果が良好な場合に前記特定点をもって前記代表点の対応点と決めることを特徴とする二次元表示画像生成方法。

【請求項14】 請求項9に記載の二次元表示画像生成方法において、前記他の映像フレームにおける特定点近傍の映像領域と、前記基準となる映像フレームの代表点近傍の映像領域間で映像の類似性を評価し、前記特定点間の相対位置の妥当性を評価し、これら両評価の結果が良好な場合に前記特定点をもって前記代表点の対応点と決めることを特徴とする二次元表示画像生成方法。

【請求項15】 請求項14に記載の二次元表示画像生成方法において、前記各評価の結果を数値化して統合し、前記対応点の位置を変化させながら前記数値を再計算し、繰り返し計算を通して前記対応点の位置精度を改善することを特徴とする二次元表示画像生成方法。

【請求項16】 請求項15に記載の二次元表示画像生成方法において、一旦全対応点の位置を固定した上で、1つの対応点のみ

を移動させながら前記各評価の結果が最良となる点を探
索し、
探索された最良点の位置を前記1つの対応点の新たな位
置とし、
これらの探索および位置の変更を順次全対応点について
行うことを特徴とする二次元表示画像生成方法。

【請求項17】 請求項16に記載の二次元表示画像生
成方法において、
前記探索および位置の変更を全対応点について行った
後、統合された各評価の数値が極値をとる条件を表すオ
イラー方程式を数値的に解くことにより、前記対応点の
位置精度を改善することを特徴とする二次元表示画像生
成方法。

【請求項18】 請求項13～17のいずれかに記載の
二次元表示画像生成方法において、
前記類似性の評価がブロックマッチングによって行われ
ることを特徴とする二次元表示画像生成方法。

【請求項19】 請求項18に記載の二次元表示画像生
成方法において、
前記ブロックマッチングでは、比較すべき映像領域の間
で色濃度の n 乗誤差総和(n は1または2)が計算され
ることを特徴とする二次元表示画像生成方法。

【請求項20】 請求項19に記載の二次元表示画像生
成方法において、
前記色濃度に対して、所定の色偏向定数を考慮して前記
ブロックマッチングを行うことを特徴とする二次元表示
画像生成方法。

【請求項21】 請求項20に記載の二次元表示画像生
成方法において、
前記色偏向定数は、前記 n 乗誤差総和が最小となるよう
決められることを特徴とする二次元表示画像生成方法。

【請求項22】 請求項21に記載の二次元表示画像生
成方法において、
前記色偏向定数は、前記映像領域間の各画素の色濃度差
の平均値であり、
前記ブロックマッチングは、前記映像領域間で各画素の
色濃度差から前記色偏向定数を減算した後に2乗誤差総
和を算出することを特徴とする二次元表示画像生成方
法。

【請求項23】 請求項13～17のいずれかに記載の
二次元表示画像生成方法において、
映像に含まれる点を特徴点と非特徴点に分類し、特徴点
である代表点について優先的に対応点を決めることを特
徴とする二次元表示画像生成方法。

【請求項24】 請求項23に記載の二次元表示画像生
成方法において、
前記特徴点である対応点を補間して前記非特徴点である
対応点を決めることを特徴とする二次元表示画像生成方
法。

【請求項25】 請求項23に記載の二次元表示画像生

成方法において、

前記基準となる映像フレームと前記他の映像フレームの
間でブロックマッチングを行い、その結果、代表点との
対応が良好である対応点を前記特徴点とすることを特徴
とする二次元表示画像生成方法。

【請求項26】 請求項23に記載の二次元表示画像生
成方法において、
前記特徴点は、撮影時刻の異なる複数の映像フレームに
おいて位置が安定的に変化している点であることを特徴
とする二次元表示画像生成方法。

【請求項27】 請求項23に記載の二次元表示画像生
成方法において、
前記特徴点は、同時に撮影された映像フレーム間におけ
る位置の変位が、その近傍時刻において同時に撮影され
た映像フレーム間においてもほぼ一定であるか、または
一定変化する点であることを特徴とする二次元表示画像
生成方法。

【請求項28】 請求項1～27のいずれかに記載の二
次元表示画像生成方法において、
前記奥行き情報に従い、映像の視点を仮想的に変更した
ときに得られるべき視点変更映像を生成することを特徴
とする二次元表示画像生成方法。

【請求項29】 請求項28に記載の二次元表示画像生
成方法において、
想定される視点の変更に伴う映像各部位の位置の変位を
前記奥行き情報から逆算し、その位置の変位に従って映
像を再構築することにより、前記視点変更映像を生成す
ることを特徴とする二次元表示画像生成方法。

【請求項30】 請求項28、29のいずれかに記載の
二次元表示画像生成方法において、元の映像が2眼カメ
ラによって撮影されたものであるとき、
これら2眼から十分に近い位置に仮想カメラを想定し、
この仮想カメラから撮影されるべき映像を視点変更映像
として生成し、
この視点変更映像と前記2眼カメラによって撮影された
現実の映像とによって多視点映像を生成することを特徴
とする二次元表示画像生成方法。

【請求項31】 請求項28～30のいずれかに記載の
二次元表示画像生成方法において、
映像の視点を、ある映像フレームを撮影した視点から別
の映像フレームを撮影した視点まで仮想的に移動させな
がら、移動経路上の任意の個所を視点とする視点変更映
像を生成することを特徴とする二次元表示画像生成方
法。

【請求項32】 請求項1～31のいずれかに記載の二
次元表示画像生成方法において、
前記奥行き情報を基準とし、映像の一部領域の大きさを
変更した上で画像を生成することを特徴とする二次元表
示画像生成方法。

【請求項33】 請求項32に記載の二次元表示画像生

成方法において、
前記映像の一部領域として奥行きが小さな領域を選定し、該領域を拡大した上で画像を生成することを特徴とする二次元表示画像生成方法。

【請求項3.4】 請求項3.2に記載の二次元表示画像生成方法において、

前記映像の一部領域として奥行きが大きな領域を選定し、該領域を縮小した上で画像を生成することを特徴とする二次元表示画像生成方法。

【請求項3.5】 請求項2.8～3.4のいずれかに記載の二次元表示画像生成方法において、
画像の処理に伴って画像枠に現れる段差を修正することを特徴とする二次元表示画像生成方法。

【請求項3.6】 請求項1～3.5のいずれかに記載の二次元表示画像生成方法において、
前記奥行き情報に従い、所望の映像領域を切り出して画像を生成することを特徴とする二次元表示画像生成方法。

【請求項3.7】 請求項3.6に記載の二次元表示画像生成方法において、
映像各部位のうち所定範囲の奥行きを持つ部位を選定することによって前記切り出しを行うことを特徴とする二次元表示画像生成方法。

【請求項3.8】 請求項3.6、3.7のいずれかに記載の二次元表示画像生成方法において、
前記切り出しの後、切り出された映像領域を別の映像に重ねることによって新たな画像を生成することを特徴とする二次元表示画像生成方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は二次元表示用の画像を生成する方法、特に、入力された単眼映像または多眼映像に加工処理を施すことによって、新たな画像を生成する方法に関する。この発明は主に、画像の視点変更、画像の一部拡大または縮小、画像領域の切り出しなどに使用可能である。

【0002】

【従来の技術】映像を撮影するカメラの数という観点からすれば、映像は単眼映像と多眼映像に大別される。単眼映像は1つのカメラ、多眼映像は複数のカメラから被写体を撮影した場合に得られる。後者はステレオ映像とも呼ばれる。

【0003】一方、この映像に所定の処理を加えて画像を出力する場合、その出力は、表示形式の観点から二次元表示画像と三次元表示画像に大別される。三次元表示画像とは、疑似立体表示のために左右両眼に与えられる画像で、静止画像の場合、通常は二枚一組のフレームからなる。動画像の場合はこのフレームの組が次々に観視者に与えられる。二次元表示画像とは三次元表示画像以外の通常の画像をいう。すなわち、二次元表示静止画像

は一枚の写真のような画像、二次元表示動画像は通常のテレビ画像のような画像をいう。以降、本明細書において「映像」といえば、主に画像処理の対象となる入力（ただし後述の視点変更映像を除く）、「画像」といえば主に処理後の出力を指す。

【0004】以上の入出力のタイプによる場合分けのうち、本発明は出力として二次元表示画像を想定する。本発明の想定する入出力の組合せは以下の2通りである。

【0005】

(1) 入力が単眼映像、出力が二次元表示画像

(2) 入力が多眼映像、出力が二次元表示画像

これらのうち、(1)では入力映像をそのまま出力することが一般的である。(2)では、多眼映像のうちの一方を選択してそのまま出力する場合が多い。多眼映像を撮影するメリットは、主に三次元表示画像の生成において発揮されるためである。本発明はこれら(1)(2)の分野において全く新しい画像処理方法を提供するものであり、対応する従来技術はないといつてよい。

【0006】

【発明が解決しようとする課題】本発明の目的は、入力映像に対して以下の処理を施し、以下の出力画像を生成することにある。この際、カメラ位置（すなわち観視者の視点であり、以下単に「視点」という）と被写体との距離の情報（以降単に「奥行き情報」という）を含む三次元位置情報の導出と利用を通して一連の処理を行う点に特徴がある。

【0007】1. 視点変更

視点を仮想的に変更したとき、映像は変化するはずである。本発明は、現実のカメラ位置を固定したまま、視点変更後に得られるべき映像（以下「視点変更映像」という）を自動生成する。

【0008】2. 画像の一部拡大・縮小

画像処理において画像の一部拡大・縮小はしばしば行われるが、本発明は奥行き情報を利用することにより、最も自然かつ効果的な画像を自動生成する。

【0009】3. 画像領域の切り出し

所望の画像領域を切り出すとき、前提として画像領域の認識が必要となる。画像領域の認識には、クラスタリング手法等いくつかの技術が提案されているが、大きな成功に至っていない。本発明は奥行き情報という全く異なる観点から、高い精度で領域の切り出しを行う。

【0010】

【課題を解決するための手段】本発明は、処理の対象となる元の映像から該映像の奥行き情報を抽出し、この情報に従って二次元表示用の画像を生成する。ここで「元の映像」とは、単眼で撮影されたものでもよいし、多眼で撮影されたステレオ映像でもよい。また、カメラで撮影された映像だけでなく、例えばアニメーション映像やコンピュータ・グラフィックスでもよい。

【0011】このとき本発明は、複数の映像フレーム間

における映像各部位の二次元的な位置の変位を検出し、この情報から前記奥行き情報を抽出する。「映像各部位」とは、映像を構成する各領域や各画素をいう。「映像フレーム」とは映像の処理単位で、通常の画像フレームの他、画像フィールド、MPEGのピクチャ等を含む。「複数の映像フレーム」は、単眼の場合は異なる時刻に撮影された複数フレーム（以下「異時刻フレーム」）をいう。多眼の場合は、同時に撮影された複数フレーム（以下「同時刻フレーム」）でもよいし、多眼カメラを構成する1つのカメラによる異時刻フレームでもよい。「二次元的な」とは、映像フレームが平面映像であることを意味する。多眼の場合でも、各カメラの映像は平面映像である。「二次元的な位置の変位」とは、平面上における位置の変位をいう。異時刻フレームなら時間経過に伴う位置の変化（すなわち動き）、同時刻フレームなら複数フレーム間の位置のずれを指す。

【0012】本発明のある態様では、前記複数の映像フレームは異時刻フレームであり、前記二次元的な位置の変位量に基づいてこれらの異時刻フレームが選択される。変位量が大き過ぎる場合は、例えば時刻差を少なくして複数フレームを選択する方法が考えられる。

【0013】本発明では、前記映像各部位の二次元的な位置の変位量を統計的に処理し、その処理結果に基づいて複数の映像フレームを選択してもよい。統計的な処理として、映像各部位の動きベクトルの分散を導出し、この値が所定値よりも大きくなるように複数の映像フレームを選択してもよい。分散の導出は既知の計算によって行われる。映像各部位の動きが小さすぎると計算誤差が大きくなるため、分散がある程度大きいほうが好ましい。

【0014】従って本発明のある態様では、分散が所定値よりも大きくなるような複数の映像フレームを選択することができない場合、二次元表示用画像の生成を中止する。誤差の大きさに起因して不自然な画像が生成される事態を回避するためである。

【0015】本発明は、こうして得られた映像各部位の二次元的な位置の変位から、映像各部位が現実の三次元空間において占める相対位置関係を導出する。その結果に従って前記奥行きを決定する。入力が単眼映像の場合、必ずしも絶対位置が求まらない。これは後述のスケールファクターの問題である。そこで、少なくとも相対位置を求める。

【0016】このとき、前記二次元的な位置の変位から前記映像各部位の三次元的な動きを算出し、この動きから三角測量の原理によって前記映像各部位の三次元空間における位置座標を算出し、その結果に従って前記奥行きを決定してもよい。単眼の場合はカメラの移動を仮定して三角測量の原理を適用する。多眼の場合は例えば左右映像に対してこの原理を適用する。

【0017】本発明では、複数の映像フレーム間にお

る被写体の対応を把握する必要がある。そこで、基準となる映像フレーム（以下「基準フレーム」という）に代表点を設定し、他の映像フレーム（以下「対象フレーム」という）における前記代表点の対応点を求め、これら代表点と対応点の位置関係を求めて、前記映像各部位の二次元的な位置の変位を認識してもよい。これは一種の画像認識技術である。ここでいう「対応点」とは、代表点に対応する点である。対応点には、「真の対応点」と「計算上の対応点」の2つの概念が存在する。対応点は本来、各代表点に対して一意的に存在するはずであり、その存在個所以外の点が対応点であるという状態は考えられない。この理想的な対応点が真の対応点である。

【0018】一方、画像処理計算から決まる対応点は、真の対応点に一致するとは限らない。これが計算上の対応点である。計算上の対応点は、真の対応点の存在個所以外に存在することがありうるだけでなく、その位置が適宜変更されることもある。後者は、例えば後述の対応点精度の改善処理を行う場合に発生する。

【0019】なお本明細書では、特に必要でない限り、「対応点」ということばを「真の対応点」と「計算上の対応点」の2つの概念に対して区別することなく用いる。

【0020】ここで、基準フレームと対象フレームが異時刻フレームであるとき、代表点と対応点の位置関係から、さらに別の対象フレームの対応点の位置を予想し、対応点を探索すべき領域を限定してもよい。最も簡単な例では、基準フレームと対象フレーム間で検出された被写体の動きが同じ速度で維持されると仮定して、別の対象フレームにおける被写体の存在個所を予想すればよい。

【0021】このとき、前記代表点のうち幾何的に特徴のある映像領域に関連するものについては、それら代表点の対応点の関連する映像領域についても前記幾何的な特徴が維持されるよう、前記対応点の位置を調整してもよい。「幾何的な特徴」を有する映像部分の例は、家の屋根の角や地平線などである。幾何的な特徴は被写体の幾何的特徴のみならず、映像中の色の変化点で構成されることもある。幾何的な特徴を維持すれば対応関係がより正確に把握されるという経験則に基づく。

【0022】このとき、幾何的に特徴のある映像領域として直線を含む領域を採用してもよい。基準フレームで直線をなした部分は対象フレームでも直線をなすと仮定するのである。対応点の位置を微調整して直線上に乗せる計算負荷は比較的軽い。

【0023】対象フレームにおける特定点近傍の映像領域と、基準フレームの代表点近傍の映像領域間で映像の類似性を評価し、この評価結果が良好な場合に前記特定点をもって前記代表点の対応点と決めてもよい。対応関係の把握に映像の類似性を採用するのである。「類似

10

20

30

40

50

性」の判断基準は、色の近さ、形の近さなどである。

【0024】ここで「特定点」とは、評価に投入される対応点の候補をいう。候補の決め方は問わない。例えば等間隔に複数の特定点をとり、これらのうちで評価結果が最良だった点を対応点とすればよい。

【0025】さらに、この類似性だけでなく、特定点間の相対位置の妥当性を評価し、これら両評価の結果が良好な場合に前記特定点をもって前記代表点の対応点と決めてもよい。例えば、基準フレームで点Aよりも右にあった点Bが、類似性評価の結果、対象フレームにおいて点Aよりも左に存在するとされた場合、その周囲の対応関係にもこうした逆転現象があれば、点Bの位置は妥当と考えることができる。別の例としては、基準フレームに取られたある領域の重心は対象フレームでもその領域の重心付近にくと仮定し、位置の評価を行ってもよい。

【0026】ここでは、2つの評価を終えてから対応点を決めてもよいし、いずれか一方が終わった段階で対応点を仮りに決め、他方の評価によってその妥当性を判断してもよい。

【0027】こうした2つの評価を経て対応点が決まった場合、本発明のある態様では、各評価の結果を数値化して統合し、対応点の位置を変化させながらその数値を再計算し、繰り返し計算を通して対応点の位置精度を改善してもよい。つまり、2種類の評価結果を重み付けして加算し、この加算結果を最適化する対応点を求める。

【0028】このときさらに、一旦全対応点の位置を固定した上で、1つの対応点のみを移動させながら前記各評価の結果が最良となる点を探索し、探索された最良点の位置を前記1つの対応点の新たな位置とし、これらの探索および位置の変更を順次全対応点について行ってもよい。

【0029】ある対応点Xを移動するとき、他の対応点を固定した上で2つの評価を行い、その結果が最良となる点をもって対応点Xの新たな位置とする。これで対応点Xの1回の位置改善が終わる。つづいて、次の対応点について同様の探索と移動処理を行う。このときも他の点は固定しておく。処理は全対応点について行い、必要に応じて二巡以上繰り返す。以下、この改善手法を「固定探索手法」と呼ぶことにする。

【0030】固定探索処理が終了したのち、オイラー方程式を用いて対応点の位置精度を高めてもよい。オイラー方程式は2つの評価の統合後の数値が極大値または極小値をとる条件を表す。この数値には、全対応点の評価結果が反映されている。オイラー方程式によれば、画像の最小単位（例えば画素）よりも高い精度で対応点の位置を改善することも可能である。

【0031】本発明では、前記類似性評価をブロックマッチングによって行ってもよい。この方法によれば、色の類似性と形状の類似性の両方が同時に評価される。

【0032】より具体的には、このブロックマッチングは、比較すべき映像領域の間で色濃度の n 乗誤差総和を計算すればよい。 n は1または2とする。 $n=1$ のときは色濃度の差の絶対値の総和を求める。

【0033】ここで色濃度に対し、所定の色偏向定数を考慮してブロックマッチングを行ってもよい（以下「バイアスド・ブロックマッチング」と呼ぶ）。同時刻フレームの場合、複数のカメラの特性によって色濃度に一定の偏位が生じやすい。同じカメラによる異時刻フレームの場合にも、時刻によって天候が変わる（つまり映像の明度等が変わる）ようなとき、同様の現象が起こる。これを打ち消すように色偏向定数を与える。

【0034】ここで色偏向定数は、 n 乗誤差総和が最小となるよう決められる。そのように決めることで色濃度の偏位が最もよく解消されると考えられるためである。

【0035】具体的には、色偏向定数は映像領域間の各画素の色濃度差の平均値とすればよい。このとき、最も一般的な2乗誤差総和が最小になるためである。前記ブロックマッチングは、映像領域間で各画素の色濃度差から色偏向定数を減算した後に2乗誤差総和を算出すればよい。

【0036】本発明のある態様では、映像に含まれる点を特徴点と非特徴点に分類し、特徴点である代表点について優先的に対応点を決める。特徴点である対応点を補間して非特徴点である対応点を決めることもできる。基準フレームと対象フレームの間でブロックマッチングを行い、その結果、代表点との対応が良好な対応点を特徴点と決めてもよい。ブロックマッチングはバイアスド・ブロックマッチングでもよい。

【0037】本発明では、異時刻フレームにおいて位置が安定的に変化している点を特徴点としてもよい。別の言い方をすれば、動きベクトルが安定している点を特徴点とする。このような点は的確に追跡されていると考えられるためである。

【0038】同時刻フレーム間における位置の変位（すなわち位置のずれ）が、その近傍時刻において撮影された同時刻フレーム間においてもほぼ一定であるか、または一定変化する点を特徴点としてもよい。このような点も追跡の精度が高いためである。

【0039】本発明の別の態様では、奥行き情報に従って視点変更映像を生成する。このために、想定される視点の変更に伴う映像各部位の位置の変位を前記奥行き情報から逆算し、前記位置の変位に従って映像を再構築する。これは、一種の画像加工技術である。例えば視点の高さを変えてみる場合、カメラの移動距離と前記奥行き情報から被写体（映像各部位）の変位（並進量と回転量）が計算できる。この計算結果に従って所期の映像を構築すればよい。

【0040】元の映像が2眼カメラによって撮影されたものであるとき、これら2眼から十分に近い位置に仮想

カメラを想定し、この仮想カメラから撮影されるべき映像を視点変更映像として生成し、この視点変更映像と前記2眼カメラによって撮影された現実の映像とによって多視点映像を生成してもよい。一般に多視点映像を得る場合、3眼以上のカメラによって映像を撮影しなければならない。そこでこの態様では、もともと2眼カメラによる映像が存在する場合、仮想カメラで撮影される映像を視点変更映像として生成し、これら3以上の映像によって多視点映像を生成する。

【0041】別の態様として、映像の視点を、ある映像フレームを撮影した視点から別の映像フレームを撮影した視点まで仮想的に移動させながら、移動経路上の任意の個所を視点とする視点変更映像を生成するしてもよい。例えば、映像フレーム1を撮影した視点を視点1、別の映像フレーム2を撮影した視点を視点2と定義すると、本発明は映像の視点を視点1から視点2に仮想的に移動させながら、移動中の任意の視点で撮影されるべき映像を視点変更映像として生成する。映像フレーム1と2が異時刻フレームを構成する場合、視点変更映像を多めに生成すれば、映像フレーム1と2の間を滑らかに変化する映像を得ることができる。これは、例えば滑らかなスローモーション映像を得る場合などに応用することができる。当然ながら、映像フレーム1、2は同時刻フレームであってもよい。

【0042】また別の態様によれば、奥行き情報を基準として映像の一部領域の大きさを変更した上で画像が生成される。映像の一部領域として奥行きが小さな領域を選定し、該領域を拡大した上で画像を生成してもよい。一般に視点から近い被写体をさらに近くに見せることにより、強調表示が可能となる。そこで小さな奥行きを持つ領域を選定し、これを拡大する。

【0043】逆の態様として、奥行きが大きな領域を選定し、該領域を縮小した上で画像を生成してもよい。もともと視線から遠い被写体をさらに遠くに見せることにより、こうした領域の重要度を落として表示するためである。

【0044】なお、こうして異なる視点からの映像や、拡大・縮小を行った後の映像については、映像領域間の遮蔽関係を考慮してもよい。例えば、視点が移動することによって奥行きの大きな被写体が奥行きの小さな被写体に遮蔽されたり、ある被写体の拡大によって背景が遮蔽される場合がある。こうした場合は、遮蔽される部分の画像を削除して画像を生成することにより、さらに自然な画像を得ることが可能となる。

【0045】視点変更映像や一部領域の拡大縮小等の画像処理に伴い、画像枠に望ましくない段差が現れることがある。本発明ではこの段差を修正してもよい。例えば、画像処理の結果、画像端部が画像中央方向に移動するような変位を受ける場合、その部分が中央方向に凹んで見えるためである。この部分は補間等の方法によって

もとの映像同様に滑らかな端部を再生すればよい。

【0046】さらに別の態様では、奥行き情報に従い、所望の映像領域を切り出して画像を生成する。切り出しは例えば、映像各部位のうち所定範囲の奥行きを持つ部位を選定することによって行われる。例えば、奥行きが10メートル以内の映像部位のみを選定する場合、視点から10メートル以内の被写体のみが切り出される。これは一種の画像認識技術である。

【0047】このときさらに、切り出された映像領域を別の映像に重ねることによって新たな画像を生成してもよい。上記の例でいえば、10メートル以内に存在する被写体を切り出し、これを別途用意した風景、同時に撮影された別の風景、同時に生成したCG映像などに張り付けて仮想的な画像を得ることができる。

【0048】

【発明の実施の形態】

実施形態1. 本発明の好適な実施形態を適宜図面を参照しながら説明する。実施形態1と2では、単眼で撮影された映像が入力されるものとする。多眼（ステレオ映像）の場合は実施形態3で説明する。

【0049】図1は本実施形態によって二次元表示画像を生成するための主な工程を示す図である。図に示すように本実施形態は以下の工程に従って、元となる入力映像から二次元表示用の画像を生成する。ここで工程1～3が映像の解析による奥行き情報の導出に関するものであり、工程4が所期の画像の生成に関するものである。まず各工程の概略を説明する。

【0050】〔工程1〕二次元動き情報の抽出

映像に含まれる被写体の動き情報を抽出する。動き情報を二次元情報であり、表示画面を座標平面上にとり、被写体のこの画面上の動きを二次元座標で記述する。

【0051】本工程では、被写体の動きの把握のために映像フレーム間の対応関係を検出する。基準フレームである時刻 t の映像フレーム（以降「フレーム t 」という）に予め複数の代表点を設定しておき、別の時刻 t' の対象フレーム（以降「フレーム t' 」という）における前記各代表点の対応点を追跡する。フレーム t と t' は異時刻フレームを形成するが、これらは時間的に隣接するフレームである必要はない。本工程の特徴は、被写体の任意の方向の動きから二次元動き情報を抽出することができる点にある。

【0052】本明細書では以降「フレーム」といえば、フィールドなど広く映像構成単位全般を指すものとし、例えば525本の走査線によって構成されるテレビ受像器の1フレームや、640×480画素によって構成されるパーソナルコンピュータの一画面のみを指すものではない。また、代表点はフレーム t のみならず、フレーム t 、 t' の両方に設定してもよい。

【0053】〔工程2〕三次元動き情報の算出

被写体の二次元的な動きが判明すれば、この被写体の三

次元空間における現実の動き情報（三次元動き情報）を算出する。この際、代表点と対応点の組を多数とることにより、被写体が実際に起こす動きを並進および回転の両運動成分によって記述する。

【0054】〔工程3〕奥行き情報の獲得

被写体の現実の動きの様子がわかれば、各時刻における被写体の相対位置関係が判明する。この関係がわかれば被写体またはその各部位（以降単に「映像各部位」ともいう）の奥行き情報が求められる。

【0055】〔工程4〕画像の生成

奥行き情報を含む三次元情報に応じて所期の画像を自動生成する。ここでは、視点変更映像、画像の一部拡大・縮小によって生成される画像、ある画像領域を切り出して生成される画像を説明する。

【0056】以上が概略である。以降、各工程を詳述する。

【0057】〔工程1〕二次元動き情報の抽出

図2は映像フレーム間の対応関係を検出するためのフローチャートである。同図に示す各ステップを説明する。

【0058】（S10）フレームtに代表点を設定

図3に示すように、まず基準フレームtに代表点を設定する。同図では、フレームtを8×8画素ごとにメッシュによって分割し、各代表点を格子点に置いている。ここでは左からi番目、上からj番目の代表点をPt

(i, j)と記述し、Pt(i, j)に対する時刻t'の対応点をPt'(i, j)と記述する。また、必要に応じてPt(i, j)のx, y座標をそれぞれ、Pt(i, j)x, Pt(i, j)yと記述する。

【0059】なお本ステップにおいて、代表点は格子点に限らず任意の箇所に配置してもよい。極端な場合は全画素を代表点とすることもできる。

【0060】（S11）対応点候補領域の設定

例えば図3に示すPt(6, 4)という代表点を考えた場合、Pt'(6, 4)の存在しうる領域を予め設定する。これは映像の動きがある限度以上に急激な場合を除き、Pt'(6, 4)がPt(6, 4)の近傍にあるという仮定に基づく。本実施形態では、例えばPt(6, 4)の近傍100×60画素の領域にPt'(6, 4)が入るものとし、Pt'(6, 4)の検出の際の計算量を削減する。

【0061】なお、本ステップについては以下の応用が可能である。

【0062】1. 映像が比較的激しく動いているときは

$$E1 = \sum \sum \{ I_t(P_t(i, j)x+u, P_t(i, j)y+v) - I_{t'}(P_{t'}(i, j)x+u, P_{t'}(i, j)y+v) \}^2 \quad (式1)$$

によって求められる。ここで2つのΣはu, vに関する。これらはそれぞれ、

u = -1, 0, 1

v = -1, 0, 1

の値を取り、仮のPt'(i, j)に対して合計9画素に

フレームt, t'が隣接するようにt'を決める。代表点の位置変化は最小に抑制され、対応点が前記領域に入らないおそれも最小となる。ただし当然ながら、予め候補領域を画面全体とする方法も考えられる。その場合、計算量が増加するが、映像の大きな動きによって対応点を取り逃すおそれは減る。

【0063】2. 本実施形態では単にPt'(6, 4)がPt(6, 4)の近傍にあると仮定したが、いったんPt(6, 4)の複数フレームにおける移動軌跡が判明したときは、この軌跡の延長線上に候補領域を決めることも可能である。映像の動きがある程度一定している場合、この方法による対応点候補領域の絞り込みは非常に有効である。

【0064】（S12）対応点候補領域における非類似度の計算

つづいて、候補領域から対応点の位置を具体的に決めていく。ただし本ステップでは、前ステップとは逆に、映像の動きが緩慢過ぎる場合に問題が生じる。動きが少ないと動き情報の抽出が困難になり、抽出された情報が大きな誤差を含む可能性があるためである。

【0065】そこでこのようなときには、予めフレームt, t'がある程度離れるようt'を決める。この際、映像各部位の変化量を統計的に処理し、例えば変化量の分散が所定値を超えるようなt'を決めればよい。なお、仮に条件を満たすt'が見つからなければ、二次元表示用画像の生成を中止する。このときは入力された映像をそのまま出力してもよい。

【0066】本ステップでは、対応点の位置を決めるべく、フレームt, t'間でブロックマッチングによる非類似度の計算を行う。前記対応点候補領域内のある点を中心とする近傍ブロックと代表点の近傍ブロックとの間で濃度の2乗誤差の総和（非類似度）を求め、これが最小になる点を計算上の対応点と決める。

【0067】図4はブロックマッチングの様子を示す図で、本実施形態では9画素で1ブロックと定義し、中央の画素を代表点とする。同図において、まずフレームt上にPt(i, j)を含むブロック1を取り、フレームt'側では対応点の候補Pt'(i, j)を含むブロック2を取る。ここで一般に、画素(x, y)の時刻tにおける画素値をIt(x, y)と記述すれば、非類似度(E1とする)は、

【数1】

関する2乗誤差総和が計算できる。そこでPt'(i, j)を候補領域内で少しずつ動かしていき、E1が最小になる点を対応点と決めればよい。

【0068】図5はPt'(i, j)ごとにE1の値を縦方向に表した模式図であり、同図の場合、非類似度が急

峻なピークをとる点Qが対応点と決まる。以降、同様に他の代表点についても対応点を決めていく。

【0069】なお、本ステップには以下の応用または変形がある。

【0070】1. ここでは濃淡画像を前提に、濃度の2乗誤差を計算したが、カラー画像の場合、RGB各濃度の2乗誤差の総和、すなわち $E1_R + E1_G + E1_B$ を非類似度としてもよい。これは他の色空間による濃度、例えばHVC濃度であってもよい。2乗誤差の代わりに単なる誤差の絶対値（残差の総和）を採用してもよい。
2. 本ステップでは1ブロックを9画素としたが、通常はある程度多数の画素によってブロックを定義することが望ましい。例えば通常のパーソナルコンピュータまたはワークステーションの高解像度画面を想定する場合、実験の結果、 16×16 画素程度のブロックによって良好な結果が得られている。

【0071】(S13) 対応点の初期位置の決定
前ステップにより、一応は対応点を求めることができた。しかしこの段階では対応点の位置が最適とは限らない。被写体の境界やエッジについては対応点が比較的正確に求まるものの、例えば変化の乏しい映像部位では、対応点の位置は相当誤差を含んでいるとみるべきである。図5でいえばE1が明確なピークをとらない状態といつてよい。図6は前ステップの結果求められた対応点と代表点の関係を示す図で、同図に示すごとく、家や木、特にそれらの輪郭等の特徴的な点については対応点

が良好に求まるが、空や地面については誤差が大きい。
【0072】そこで本ステップと次ステップにより、対応点の位置補正を行う。本ステップでは、まず対応点の初期位置という概念を導入し、実際にその初期位置を決める。次ステップで繰り返し計算による位置精度の改善を図る。本ステップにおける対応点初期位置の決定には以下の方針が考えられる。

【0073】1. 前ステップで求められた全対応点を同等に扱う

全対応点の位置をそのままそれらの初期位置として次ステップへ投入する。

【0074】2. 対応点の扱いに差異を設ける

当初よりある程度正しい位置にあると思われる対応点（以降「特徴点」という）の位置はそのまま初期位置として使い、そうでない対応点（以降「非特徴点」という）の初期位置については、前記特徴点のそれをもとに決めるというものである。ここで特徴点としては以下の点を想定することができるが、これらの点は現実には一致することもあり。

【0075】(1) 前ステップのE1が明確なピークを示した対応点

このような対応点の位置精度は一般に高いためである。

$$E2 = \{Pt'(i,j)x - St'(i,j)x\}^2 + \{Pt'(i,j)y - St'(i,j)y\}^2 \quad (式2)$$

が相対位置の妥当性評価式となる。この式だけを考えれば、E2が最小値になるときに

【0076】(2) 直交エッジ成分が多く存在する箇所にある対応点

建物の角などの部分では対応点の位置がかなり正しいと思われるためである。

(3) フレーム t, t', \dots において位置が安定的に変化している対応点

ここで変化の安定性は、動きベクトルの一定性といえることができる。ここでは、フレームの進行に伴って移動方向、移動距離が一定な対応点を選ぶ。具体的には、例えば動きベクトルのばらつきが所定値以下の対応点を選定する。こうした対応点は的確に追跡されているはずであり、代表点と正しい対応関係にあると判断できるためである。ただし、例えば映像を撮影するカメラが不規則に動いた場合には、その影響を加味して判断する。

【0077】こうして特徴点を選定されれば、これらはそのまま初期位置として使い、非特徴点の初期位置は特徴点の位置を補間したり、または特徴点の近傍から順に決めていくことができる。すなわち、前ステップによる非特徴点の位置精度は低いため、精度の高い特徴点から幾何的に非特徴点の初期位置を与えるのである。なお当然ながら、(3)の特徴点を見つける場合にも前ステップの方法を有効利用できる。

【0078】以上、特徴点の選定に基づいた対応点初期位置の決定方法を述べたが、その他、動的計画法（ダイナミック・プログラミング）を用いて対応点の初期値を求めてもよい。

【0079】(S14) 対応点の改善処理

対応点の位置の妥当性を評価するために式を導入し、繰り返し計算によって位置精度を改善する。S12ステップでは非類似度を評価する式1を導入したが、ここではさらに、対応点間の相対位置関係の妥当性を評価する式を導入し、これら2つ評価結果を統合して位置の改善を図る。

【0080】図7は相対位置を評価する原理を説明する図である。同図において、各点はそれぞれ対応点を表している。このうち、図中の $Pt'(i, j)$ を中心に考えると、これには以下の4つの対応点、

【数2】 $Pt'(i-1, j), Pt'(i+1, j), Pt'(i, j-1), Pt'(i, j+1)$ が隣接している。 $Pt'(i, j)$ は通常、これら4つの点の重心付近に存在すると考えるのが妥当である。これは、映像各部位が動いても画素単位の微視的な見方をすれば相対位置関係がほぼ保たれるという経験則に基づいている。なお、この性質を数学的にいえば、 (i, j) の関数 $Pt'(i, j)$ の2次微分がほぼ0であるということにほかならない。

【0081】従って上記4点の重心を $(St'(i, j)x, St'(i, j)y)$ と表記すれば、

【数3】

ば、E2が最小値になるときに対応点の位置が最も望ま

しい状態となる。

【0082】本ステップでは、式1および式2の評価結果を適当な結合定数kで加算し、

$$E = E1 / N + k \cdot E2 \quad (\text{式3})$$

で表されるEを最終的な評価式とする(Nはブロックマッチングの際に定義された1つのブロックに含まれる画素数である)。すなわち、まず各対応点についてEを計算し、続いて全対応点のEの総和ΣEを計算し、ΣEが最小値となるよう、各対応点の位置をすこしずつ変化させる。ΣEの値が収束するか、または繰り返し計算を所定の上限回数に達するまで行い、改善処理を施す。より具体的には、各対応点の位置を変化させるとき、以下のいずれかの方法を実施すればよい。

【0083】(1) オイラー方程式を解く方法

ΣEが極値(ここでは極小値)をとる条件を示すオイラー方程式を数値的に解くことによって対応点を得る。この手法自体は既知である。これは、各代表点を含むブロックでの画像傾き情報と、対応ブロック間の画素差分情報から改善すべき方向を見出し、これに基づいて対応点の位置を初期位置から徐々に動かしていき、最終解を

【0084】(2) 固定探索手法

まず、対応点候補領域において、改善対象の対応点のEが最小になる点を探し、これを新たな対応点とする。このとき、他の点の位置を不動とみなして探索を行う点に特徴がある。この処理を順次全対応点に対して行う。

【0085】(3) 混合手法

(2)の手法によれば、画素単位の精度で対応点の位置が求まる。一方、(1)によれば、理論上画素単位以下の精度で位置を求めることができる。そこで、まず

(2)の手法によって画素単位の精度で対応関係を求め、しかる後にオイラー方程式を適用して精度を高めることも可能である。

【0086】なお実験によれば、同じ精度で比較した場合、(2)の手法によって(1)よりも短時間で好ましい解が得られている。

【0087】図8は図6の対応点候補に対して本ステップの改善処理を行った結果を示す図である。実験によれば、カラー画像の場合、kは5〜200程度で良好な結果が得られることがわかった。図6と図8はともに模式的な図であるが、実験の結果、実際にこれらの図に近い改善が見られた。

【0088】以上が工程1の詳細である。なお本工程には、以下の応用または変形がある。

1. E2導出の際、上下左右の4点のみならず、斜め方向の4点を加えた計8点の重心を考える。いかなる組合せが最適であるかは映像の種類にも依存するため、適宜実験によって決めていくことが望ましい。

【0089】2. 式3による評価は、E1のみによる評価結果が思わしくなかった対応点から優先的に行う。こ

れはE1の結果が悪い対応点は一般に位置の誤差が大きいと考えられるためであり、こうした対応点の位置を早期に、かつ大幅に改善することが望ましいためである。

【0090】3. 位置改善の際、幾何情報も利用する。フレームtにおいて幾何的に特徴のある領域、例えば直線を形成していた複数の代表点については、それらの対応点も直線を形成するように位置を補正する。これは映像上直線に見える部分は現実の三次元空間でも直線である可能性が高く、一方、三次元空間の直線はフレームt'でも直線となるべきだからである。本来奥行きは直線に沿って一様に変化するものであり、直線に沿う変化は視覚的に容易に把握されるため、この方法による改善効果は大きい。かかる改善をしない場合、最終的に生成される画像において直線に沿う奥行きにむらが生じ、不自然な画像になることがある。

【0091】4. さらに別のフレームについても対応点を求める。本工程ではフレームtに対するフレームt'の対応点を求めたが、さらに第三のフレームt''における対応点も求め、映像各部位の平均化された動きを求めることができる。この方法は、フレームt'における対応点位置を改善していくのではない。多くのフレームで対応点をとることにより、対応点の位置とそのフレームが撮影された時間から映像各部位の動きを統計的に決めていくのである。

【0092】[工程2] 三次元動き情報の算出

工程1により、映像各部位の画面上の二次元的な動きが判明した。工程2ではこの情報から各部位の三次元的な動きを算出する。映像は被写体の現実の動きを平面に投影したものであり、本工程では代表点と対応点の位置関係からもとの動きを導出する。

【0093】一般に被写体の三次元空間における動きは、並進運動と回転運動の合成として記述することができる。ここではまず、動きが並進運動のみで構成される場合の計算方法を説明し、後に一般化された方法を概説する。

【0094】1. 動きが並進運動のみの場合

図9はある点Pの画面上の移動と三次元空間での現実の移動の対応を示す図である。同図では画面上の二次元座標を大文字X等で、現実の三次元座標を小文字x等で表記するものとし、三次元座標のうちx、y軸を画面上に、z軸を奥行き方向にとっている。また、視点から画面までの距離を1とする。

【0095】この図に示す通り、P(X, Y)は画面上をP'(X', Y')へ移動するが、この間、この点は三次元空間においてS(x, y, z)からS'(x', y', z')へと移動する。ここで、

【数4】

$$(x', y', z') = (x, y, z) + (a, b, c)$$

とすれば、画面までの距離が1なので、

$$X = x / z, \quad Y = y / z$$

$$X' = x' / z', \quad Y' = y' / z'$$

となる。これを解けば、

$$X' = (Xz + a) / (z + c)$$

$$Y' = (Yz + b) / (z + c)$$

$$(a - X'c)(Y' - Y) = (b - Y'c)(X' - X) \quad (式4)$$

式4は画面上の動き情報で表現されているため、工程1で得られた情報によって未知数 a 、 b 、 c を決めることができる。しかしこの際、現実には k 倍の大きさの物体が k 倍離れたところを k 倍の速さで移動するケースにおいて、この k の値(スケールファクター)を決めることはできず、 a 、 b 、 c についてはそれらの比のみを求めることが可能となる。数学的にいえば、 (X, Y) と (X', Y') の対応を3組与えても、この連立方程式を行列表示した際の係数行列のランク(階数)は高々2であり、 a 、 b 、 c は相対値としてしか決まらない。そこで本工程では、仮に $c=1$ と正規化して a 、 b を表すことにする。比のみでも、次工程による処理が可能のためである。

【0097】並進運動の別の解法として、式4から誤差

$$a \Sigma (Y' - Y)^2 - b \Sigma (X' - X)(Y' - Y) - c \Sigma (Y' - Y)(XY' - X'Y) = 0 \quad (式9)$$

$$-a \Sigma (X' - X)(Y' - Y) + b \Sigma (X' - X)^2 + c \Sigma (X' - X)(XY' - X'Y) = 0 \quad (式10)$$

$$-a \Sigma (Y' - Y)(XY' - X'Y) + b \Sigma (X' - X)(XY' - X'Y) + c \Sigma (XY' - X'Y)^2 = 0 \quad (式11)$$

以上が並進運動に関する計算方法の例である。

【0100】2. 動きが回転運動を含む場合

回転運動は x 、 y 、 z 方向の3つの変位と各軸を中心とする3つの回転角、例えば α 、 β 、 γ によって記述することができる。回転角はオイラー角またはロールピッチ法などによって表現することができる。

【0101】ここで上記合計6つの変数を決定すればよいが、ここでも上述のごとくスケールファクターが決まらないため、ある変数を1として各変数の比を求める。理論上、代表点と対応点を5組とれば運動を記述することができる。

【0102】ここで注意すべきは、組のとりかたによっては動きの様子が線形解法によって求まらないことがある点である。こうした場合を考慮する際、組を8以上とればよいことが知られている。8組の変化から線形解法によって回転運動を記述しうる根拠については、例えば

$$(x, y, z) = R^{-1} \{ (x', y', z') - (a, b, c) \} \quad (式12)$$

をカメラの動きと考える。

【0105】図10はカメラの三次元移動とある点Pの画面上の移動から点Pの三次元座標を導く原理を説明する図である。同図からわかるように、この原理は一般に三角測量の原理として知られるもので、位置の異なる2点から点Pの方向を見たとき、点Pの現実の位置(図中の点S)はそれら2つの視線の交点に存在するというものである。

【0106】同図では、時刻 $t \sim t'$ の間にカメラが矢印で示すように式12に従って移動したとする。フレーム t では点Sが点 P_t に、 t' では点 $P_{t'}$ にそれぞれ投

となるため、 z を消去し、次式が求められる。

【0096】

【数5】

e を、

【数6】

$$e = \{ (a - X'c)(Y' - Y) - (b - Y'c)(X' - X) \}^2 \\ = \{ (Y' - Y)a - (X' - X)b - (XY' - X'Y)c \}^2 \quad (式5)$$

と定義し、代表点と対応点の全対応関係について e の総和 Σe をとり、この値を最小にする a 、 b 、 c を次の式から求めてもよい。

$$\frac{d(\Sigma e)}{da} = 0 \quad (式6)$$

$$\frac{d(\Sigma e)}{db} = 0 \quad (式7)$$

$$\frac{d(\Sigma e)}{dc} = 0 \quad (式8)$$

より具体的には、式6~8はそれぞれ次の形に展開される。

【0099】

【数7】

$$(式9)$$

$$(式10)$$

$$(式11)$$

「動きからの単眼立体視による形状認識の線形解法について」(出口・秋場、計測自動制御学会論文集vol.26, No.6, 714/720 (1990))などに示されている。

【0103】[工程3] 奥行き情報の獲得工程2によって映像各部位の三次元的な動きの相対量がわかった。工程3では、この相対量から各部位の奥行き情報を導出する。本工程では説明のために、被写体は静止しており、それを撮影するカメラの側が動くものと仮定する。映像処理の際には被写体とカメラの相対運動が問題となるため、この仮定によって良好な結果が得られる。

【0104】映像のある部位の動きを回転行列 R と並進ベクトル (a, b, c) により、

$$(x', y', z') = R(x, y, z) + (a, b, c)$$

と表す場合、この逆変換、

【数8】

影されている。点Sは図中の2つの直線 L_t 、 $L_{t'}$ の交点にある。

【0107】この条件下、カメラの方向と L_t 、 $L_{t'}$ のなす角 θ_t 、 $\theta_{t'}$ は既知であり、一方カメラの移動方向と距離が判明しているため、点Sの三次元座標を求めることが可能となる。この座標により、映像各部位の奥行き情報が判明する。

【0108】ここで注意すべきは、前述のごとく $c=1$ という正規化のため、求められた座標も一定の割合で拡大または圧縮されていることである。しかしこの場合でも、奥行き情報は一様に拡大圧縮されているため、奥行

きの相互関係は正しい。

【0109】以上が本工程の概要であるが、本工程では前工程までの誤差を考慮する必要がある。誤差により、通常は前記 L_t 、 $L_{t'}$ が計算上交わらないためである。こうした事情に配慮し、本工程では両直線の最接近

$$\begin{aligned} L_t &: (x, y, z) + \alpha (u, v, w) \\ L_{t'} &: (x', y', z') + \beta (u', v', w') \end{aligned} \quad (式13)$$

と表すことができる。従って、

【数9】

$$e = \{(x + \beta u) - (x' + \alpha u')\}^2 + \{(y + \beta v) - (y' + \alpha v')\}^2 + \{(z + \beta w) - (z' + \alpha w')\}^2 \quad (10) \quad 【数10】$$

$$(u^2 + v^2 + w^2) \alpha - (uu' + vv' + ww') \beta + (x - x')u + (y - y')v + (z - z')w = 0$$

$$(u'^2 + v'^2 + w'^2) \beta - (uu' + vv' + ww') \alpha + (x - x')u' + (y - y')v' + (z - z')w' = 0$$

を解いて α 、 β を求め、最終的に点 S の奥行き値を、

$$【数11】 \{(z + \alpha w) + (z' + \beta w')\} / 2$$

とすればよい。ここで仮に誤差が0だったとすれば、この座標は両直線の交点の z 座標に一致する。

【0111】また別の方法として、これら両直線を一旦フレーム t の画面に透視投影し、投影後の最近接点の z 座標を求めることもできる。ここで L_t は代表点である一点に投影され、一方 $L_{t'}$ は一般に直線に投影される。 $L_{t'}$ が式13で表記されるならば、投影後の各点の x 、 y 座標は、 $L_{t'}$ 上の各点のそれらをその z 座標で割ることにより、

$$x = f(x' + \beta u') / (z' + \beta w') \quad (式14)$$

$$x = (m^2 X - kn - kmY) / (k^2 + m^2) \quad (式16)$$

$$y = (k^2 Y - mn - kmX) / (k^2 + m^2)$$

となる。ここで点 T に対応するもとの $L_{t'}$ 上の点を $E(x'', y'', z'')$ とすれば、点 E は、式16を式14に代入して β を求め、これを $L_{t'}$ の式に代入することによって求められる。 β は、

$$\beta = (xz' - fx') / (fu' - xw')$$

であるため、これを式13へ代入し、点 E の z 座標 z'' は、

$$z'' = z' + w' (xz' - fx') / (fu' - xw')$$

と求まる。これを点 S の奥行き値とすればよい。

【0114】なお、画像処理の際の誤差によって上記の奥行きが負の値となる場合（点 S がカメラの後方に存在することになる場合）、この計算結果は信頼することができない。このときは正の奥行き値を持つ近傍の代表点から補間するなどの処理を行う。

【0115】以上、いずれの方法をとるかは別として、求められた映像各部位の奥行きは、例えば代表点ごとに数値として与えればよい。図11はフレーム t において各代表点に数値が与えられた状態を示す図で、例えば $P_t(2, 3)$ 、 $P_t(4, 3)$ それぞれの奥行きは100、200となり、後者の実際の位置は前者よりも2倍離れたところにあることがわかる。

【0116】〔工程4〕画像の生成

工程3で求められた奥行き情報に応じて、所期の画像を

点の中点の z 座標を点 S の奥行き値と近似する。これを数式によって説明する。

【0110】上記 L_t 、 $L_{t'}$ の方向ベクトルをそれぞれ (u, v, w) 、 (u', v', w') とする。ここで実数パラメータ α 、 β により両直線は、

とし、 e を最小にする α 、 β を $de/d\alpha = 0$ 、 $de/d\beta = 0$ より求める。すなわち、

【数10】

$$y = f(y' + \beta v') / (z' + \beta w') \quad (式15)$$

と書くことができる。ここで f は視点からフレーム t の画面までの距離で、実際には $f = 1$ などとして扱えばよい。式14、15から β を消去すれば投影後の直線（以下 L_i という）が以下のように求まる。

$$【0112】 kx + my + fn = 0$$

ただしここで、

$$k = v'z' - w'y', \quad m = w'x' - u'z', \quad n = u'y' - v'x'$$

とおいている。

【0113】求めるべき最近接点は、代表点 P_t から L_i に下ろした垂線と L_i の交点（以下 D とする）であり、その座標は、

$$(式16)$$

生成する。いままでの工程では少なくとも2つの映像フレームから必要情報を抽出したが、本工程では1つの映像フレームをもとに所期の画像を得ることができる。

【0117】(1) 視点変更映像

図12、13は元の映像と視点変更映像の対応関係を示す図である。図12には元の映像が示され、この映像は奥行きが小さい順に「木」、「家」、「人」の映像領域を含んでいる。一方、図13には視点変更映像が示されている。ここでは、視点を仮想的に右上に移動した場合を例にしている。

【0118】これらの図からわかるように、本発明によれば、実際にカメラの位置を移動しなくとも、異なる視点からの映像を得ることができる。工程3により、映像各部位について奥行き情報を含む三次元位置座標が把握されているためである。ここでは視点が右上に移動すると仮定したが、これは被写体が左下に移動すると考えることと等価である。この左下への移動は、工程3で説明したように、回転と並進によって表すことができる。そこで、本実施形態の工程1～3を逆向きに辿ることにより、この仮想的な三次元移動から画像上の二次元移動を逆算することが可能となる。この結果、図13に示す画像が得られる。工程1～4の画像生成過程には人が恣意をさし挟む余地はなく、得られる画像は極めて自然なも

のとなる。

【0119】本工程では、遮蔽関係を考慮して画像を生成することが望ましい。例えば図13の場合、視点の移動に伴って「木」の下端が「家」の屋根に隠される。そこで「木」の下端の部分を「家」の画像データで塗りつぶすことにより、自然な画像を生成する。実際のソフトウェア処理では、奥行き大きい映像部位から順に、視点移動後の画像を生成していけばよい。遮蔽関係を計算によって把握したいときは、新たな視点から映像各部位を見込む視線ベクトルが一致するか否かを判定すればよい。視線ベクトルが一致する2つの部位A、Bがある場合、部位Aのほうが部位Bよりも視点に近ければ、部位Bが部位Aによって隠されることになる。この情報に従って画像を生成してもよい。

【0120】なお、視点変更映像を生成する場合、当然ながら、視点の仮想的な移動距離が短いほうが正確な画像を得ることができる。この事実を利用して、視点変更映像には次のような用途も考えられる。

【0121】1. 2眼映像から多視点映像の生成

2眼カメラによってステレオ映像が得られているとき、仮想的に第3のカメラを設けて多視点映像を生成する。2眼カメラの間隔は一般に小さいため、例えばこれらとともに小さな三角形をなす点を視点とする視点変更映像を生成する。この映像は比較的正確であり、もとの2眼カメラによる映像と合わせることで良好な多視点映像を形成することができる。

【0122】2. スローモーション映像の生成

例えば時間的に最も近い2枚の異時刻フレームをフレーム t 、 t' とする。これらのフレームを撮影した視点をそれぞれ視点 t 、 t' とする。視点はフレーム t 、 t' の間に視点 t から視点 t' に実際に移動しているが、この間の映像は存在しない。そこで視点 t 、 t' 間に仮想的な視点を設け、視点変更映像を生成する。視点変更映像を多数設けて順次表示すればスローモーション映像が得られる。この際、次の効果がある。

【0123】a. 本来ならコマ送りの映像となるべき映像の動きが極めて滑らかになる

b. 時間的に近いフレーム間では一般に視点の移動が少ないため、非常に良好な映像が得られる

c. 視点 t 、 t' 間の視点の移動経路を変化させれば、異なるイメージのスローモーション映像が得られる。

【0124】なおこの技術は、当然ながら同時刻フレーム間に適用することもできる。

【0125】(2) 一部拡大・縮小画像

画像表示技術の1つに強調表示がある。強調表示の一例として、近い被写体はより近く、遠い被写体はより遠くに配置しなおす場合がある。画像にメリハリを付けるためである。

【0126】こうした要請に答えるべく、本発明では奥行き情報に基づいて画像の一部拡大・縮小を行う。図1

4は図12の画像の一部を拡大して生成された画像を示す。ここでは拡大処理の対象を奥行きの小さな領域、すなわち「人」としている。この結果、観視者に最も近い「人」がさらに手前に知覚され、効果的な強調表示が可能となる。ここでも(1)同様、遮蔽関係を考慮することが望ましい。

【0127】ここで注意すべきは、最小の奥行きを持つ領域を拡大する限り、原則として拡大率に制限がないことである。その領域が理論上奥行きゼロの距離に知覚されてもよいためである。一方、中間の奥行きを持つ領域、例えば図14の「家」を拡大する場合、これは「人」よりも手前に知覚されるべきではないため、自ずと制限が生じる。この制限を守らない場合、非常に不自然な画像が生成される。本発明のように、奥行き情報に基づいて拡大を行う場合、例えば、「最小の奥行きを持つ領域のみを拡大し、最大の奥行きを持つ領域のみを縮小する」という規則を設けることが可能となり、現実にはマッチした自然な画像(自然法則に則った画像)を得ることができる。

【0128】ここでは自然な画像の生成について触れたが、逆に、敢えて不自然な画像を作りたいという要望もある。例えばゲームの画像において、遠い個所にあるものを拡大して一種のちぐはぐ感を創出するような用途が考えられる。いずれにせよ、本実施形態の特徴はこうした自然感、不自然感の選択を意識的に行うことのできる点にある。従来の手法では、ある領域を拡大してみたら偶然自然な画像が得られた、というような事態が発生しがちだったが、本実施形態はこうした事態を解消する。

【0129】なお、本工程の結果自然な画像を得られたとして、この画像に対し、さらに上記(1)または後述(3)の処理を施す場合、まず拡大・縮小した領域の奥行きを変更することが望ましい。例えば、2倍の長さに拡大した領域については、その奥行きを約 $1/2$ にする。逆に、 $1/2$ の大きさに縮小した領域は奥行きを約2倍にする。これは領域の大きさと奥行きがほぼ反比例する事実による。この配慮により、後段の処理においても自然な画像を得ることができる。

【0130】また、(1)と(2)の画像処理に際して、画像枠に現れる段差を修正して最終画像を生成してもよい。例えば、図12の元の映像から図13の視点変更映像を生成する場合、一般に図13に含まれるすべての映像部位と図12のそれらが完全に対一に対応することはない。図13の右上端は、図12の場合に比べて、より「家」の右上端から遠いため、図13の右上端付近の映像の元となる映像は図12に含まれていない。従って、図12に含まれる映像情報から単純に図13の映像を生成する場合、右上端付近が欠ける。この部分が図13に示す理想的な画像枠から内側に凹む。同様の理由から、図12に含まれる剰余の映像情報は図13の画像枠から外に飛び出す。

【0131】そこでここでは、凹部に画素を充当し、一方凸部から余剰画素を切除することにより、画像枠（画像領域の外形で、ここでは矩形）をもとのまま維持するものとする。充当は、例えば近傍の画素と同じ色の画素によって行う。（2）の一部拡大・縮小の場合も、それぞれ凸凹が発生すると考えられるため、同様の処理を施す。こうした処理により、不自然な画像枠の表示を回避することができる。

【0132】（3）切り出しによる画像所期の領域を切り出して画像を生成する。例えば図12において、「人」「家」「木」の奥行きがそれぞれ3、10、20メートルであるとする。ここで仮に、「人」のみを切り出す場合、例えば「奥行き5メートル以内」という条件を与えて各部位の奥行きを検索、判定していけばよい。「家」の場合は「奥行き5～15メートル」などとする。

【0133】図15は図12の画像のうち、「家」を切り出して生成された画像を示している。所望の領域が切り出された後は、それ以外の領域を空白にしたり、別の映像に切り出した領域を貼り付けることができる。

【0134】このように、本発明は一種の画像認識・加工技術をも与える。従来、画像領域の切り出しは人手作業または色によるクラスタリング等の手法によって行われたが、本発明は奥行き情報という新しい観点から、相当精度の高い領域認識手法を提供する。

【0135】以上が実施形態1に係る二次元表示画像生成方法の説明である。本発明によれば、正確な奥行き情報の導出とその利用により、新しい画像処理技術が開示される。一連の処理は、例えばソフトウェア的にすべて自動化することができるため、本発明の用途は広範である。

【0136】実施形態2。実施形態2では実施形態1の実施に最適な装置を説明する。

【0137】図16は実施形態1を実施するためのハードウェア構成例を示す図である。

【0138】同図において、処理対象となる映像は映像入力回路20から入力され、ここでA/D変換を受ける。変換後の映像はフレームメモリ制御回路22によってフレームメモリ24に格納される。フレームメモリ24の後段には、複数の映像フレームを読み出して対応点を検出する対応点検出回路26が置かれている。この回路は実施形態1の工程1をハードウェア的に実施するもので、ブロックマッチングには、例えばMPEGのエンコーダ回路を採用することができる。

【0139】対応点検出回路26において検出された対応点の座標は、いったん対応点座標メモリ28に格納され、動き検出回路30によって適宜読み出される。動き検出回路30は実施形態1の工程2、3を実施するもので、並進・回転運動から被写体の三次元相対位置を算出する。

【0140】つづいて三次元位置情報は画像生成回路32へ入力される。この回路はフレームメモリ24から元の映像を読み出し、所定の処理を加えて画像を生成する。画像生成回路32の前段には、外部から各種指示を受け付ける指示入力部34が設けられている。

【0141】こうして生成された画像は、画像出力回路36によるD/A変換を受け、図示しないディスプレイ装置へ出力される。

【0142】以上の構成による本装置の動作を説明する。

【0143】まず、カメラによって撮影された映像またはビデオ再生装置によって再生された映像等が映像入力回路20から入力される。この映像はフレームメモリ24に格納される。フレームメモリ24から読み出された複数の映像フレームに対し、対応点検出回路26、動き検出回路30によって被写体の奥行き情報が獲得される。つづいて画像生成回路32がこの奥行き情報に従い、視点変更映像等の画像を生成する。この際、前述の指示入力部34からの入力により、視点変更、拡大・縮小、切り出しを行う。こうした指示は、既知のユーザー・インタフェースによって実現可能である。

【0144】以上、本装置をワークステーションに組み込んで映像の奥行き情報を求めた実験の結果を図面によって説明する。

【0145】図17～22は本装置による画像生成の過程を示すもので、いずれもディスプレイ上に表示された中間調画像の写真であり、約640×480画素領域で構成されている。

【0146】図17と図18はそれぞれフレームtとフレームt'の映像で、カメラの移動によって多少映像に動きがある。図19はフレームtをメッシュ分割して代表点を設けた状態を示し、図20はフレームt'における対応点の初期位置を示している。ここでは特徴点を優先的に扱い、代表点を中心とする16×16画素の領域でブロックマッチングを行った結果の最良点を初期位置とした。

【0147】図21は実施形態1の式3、つまり対応点の位置関係をも考慮して対応点位置を改善した結果を示す図で、図20に比べて位置の大幅な改善が見られる。

【0148】図22は獲得された奥行き情報を濃淡で表したもので、薄い画素ほど奥行きが小さい。同図からわかるように、奥行き情報がかなり正確に得られている。

【0149】以上、本装置によれば実施形態1を円滑に実施することができる。この際、特にブロックマッチングのソフトウェアによる実行時間を考慮すれば、ここをハードウェアで行うことによる処理速度の改善効果は大きい。

【0150】本装置を製品とする場合の態様としては、例えばパーソナルコンピュータやワークステーションに図16の構成を有するアド・オンカードを搭載する方

法、テレビ受信装置やビデオ再生装置に図16の構成を予め内蔵する方法等が有効である。本装置をカメラと組み合わせた場合、従来必要とされたレーザ、赤外線、超音波等による奥行き測定が不要となる。

【0151】実施形態3。実施形態1、2では、入力映像が単眼カメラによって撮影されるものとした。ここでは、多眼カメラによるステレオ映像を入力映像とする場合の二次元表示画像生成方法を、実施形態1との相違点を中心に説明する。

【0152】図23は実施形態3によって二次元表示画像を生成するための主な工程を示す。同図と実施形態1の図1との主な違いは以下の通りである。

【0153】1. 工程1の「動き情報」が「変位情報」に変更される

実施形態1では異時刻フレームを扱ったが、実施形態3では基本的に同時刻フレームを扱う。同時刻の場合、被写体に動きを定義することはできないため、代わりに同時刻フレーム間の被写体の位置のずれ(変位)の情報を抽出する。

【0154】2. 工程2が不要となる

図1の工程2「三次元動き情報の算出」に対応するステップがない。多眼の場合、はじめから図10の状態では撮影がなされるため、三角測量の原理によって直接的に奥行き情報が獲得できるためである。

【0155】なお、複数カメラの相対位置関係に狂いが発生する多眼カメラシステムを用いる場合、この狂いを補正するセルフキャリブレーションを行ったほうがよい。この場合、工程2はセルフキャリブレーション工程となる。セルフキャリブレーションの手法については、例えば、富田、高橋「ステレオカメラのセルフキャリブレーション」(情報処理Vol.31, No.5 (1990) 650~659ページ)、特開昭63-293038号公報、特開昭63-293039号公報などに示されている。

【0156】以下、実施形態3の工程1~3を説明する。

【0157】[工程1] 二次元変位情報の抽出

実施形態1の説明において、「動き」を「変位」に置き換える他、フレーム t 、 t' の組をフレーム1、2に置き換えればよい。フレーム1、2はそれぞれカメラ1、2から撮影された映像を指し、撮影時刻は t で固定とする。実施形態3では、最低これら2枚のフレームのみから最終画像を得ることができる。すなわち多眼撮影の場

$$E1 = \sum \sum \{ I1(P1x+u, P1y+v) - I2(P2x+u, P2y+v) \}^2 \quad (\text{式18})$$

と簡単になる。この式が濃淡画像の場合の通常のブロックマッチングを表す。

【0163】一方、バイアスド・ブロックマッチングで

$$E1 = \sum \sum \{ I1(P1x+u, P1y+v) - I2(P2x+u, P2y+v) - \alpha \}^2 \quad (\text{式19})$$

とすればよい。カラー画像の場合、 α は α_R 、 α_G または α_B であり、RGBそれぞれの画像において求めたE1の和、すなわち $E1_R + E1_G + E1_B$ でマッチング

合は、入力静止映像であってもよい。その他、実施形態1の工程1との相違は以下の通りである。

【0158】(1) 実施形態1のS11(対応点候補領域の設定)では、映像の動きの激しさまたは各部位の移動軌跡に基づき、異時刻フレームの選択または対応点候補領域を絞り込んで対応点検出処理の計算量削減を行った。実施形態3では絞り込みの方法を以下のように変更し、同様に有効な計算量削減を実現する。

【0159】まず、通常どおり多眼カメラが水平に設置されると仮定する。このとき、対応点の y 座標(上下方向の座標)はほぼ等くなる。この仮定と画像処理に伴う誤差およびカメラの設置誤差を考慮し、対応点候補領域を横長の帯状領域に限定する。さらに、フレーム t' ($t' = t - 1$)において対応する代表点の位置の差が x であれば、フレーム t における対応点探索領域も、やはり差が x となる近傍に限定することができる。

【0160】(2) 実施形態1のS12(対応点候補領域における非類似度の計算)では、映像の動きが緩慢過ぎる場合に統計処理を導入したが、実施形態3ではこの作業も不要である。

【0161】(3) 実施形態1のS12同様、実施形態3でも対応点の位置を決めるべくブロックマッチングを行うが、ここではバイアスド・ブロックマッチングを採用したほうがよい場合がある。バイアスド・ブロックマッチングは、多眼カメラを構成する各カメラが異なる特性を持つ場合、有効に機能する。例えば、カメラ2がカメラ1よりも青味がかかった映像を送出するとすれば、フレーム2の色濃度から青(B)の成分を一定量差し引いた後(すなわち所定の色偏向定数 α_B を引いた後)、ブロックマッチングを行うべきである。こうした処理を行わない場合、最良マッチングを取り逃がすおそれがある。ただし実際には、例えばRGBで色濃度を表す場合、前記の青(B)のみならず、赤(R)と緑(G)についてもそれぞれ色偏向定数 α_R 、 α_G を引くべきである。

【0162】図4および式1を基礎として、バイアスド・ブロックマッチングを式で説明する。ここでは、実施形態1で用いた $P_t(i, j)$ をフレーム1、2に対応して単に $P1$ 、 $P2$ と表記し、 $I_t(i, j)$ も同様に $I1$ 、 $I2$ と表記する。このとき式1は、

【数12】

は、式18を、

【数13】

を行う。さらに見やすさを考えて、 $I1(P1x+u, P1y+v)$ を単に $I1$ 、 $I2(P2x+u, P2y+v)$ を単に $I2$ と表記すれば、式19は、

$$E1 = \sum \sum (I1 - I2 - \alpha)^2 \quad (\text{式 } 20)$$

となる。I1、I2はu、vの関数であるが、 α は定数である。

【0164】 α の最適値を考える。カメラ1、2は同じ被写体を撮影しているはずだから、フレーム1、2の映

$$\begin{aligned} E1 &= \sum \sum \{ (I1 - I2)^2 - 2\alpha(I1 - I2) + \alpha^2 \} \\ &= \sum \sum (I1 - I2)^2 - 2\alpha \sum \sum (I1 - I2) + \sum \sum \alpha^2 \end{aligned} \quad (\text{式 } 21)$$

であり、領域の総画素数をNと仮定すれば、 $\sum \sum 1 = N$ だから式21は、

$$E1 = \sum \sum (I1 - I2)^2 - 2\alpha \sum \sum (I1 - I2) + N\alpha^2 \quad (\text{式 } 22)$$

となる。従って、

$$dE1/d\alpha = -2\sum \sum (I1 - I2) + 2N\alpha$$

であるから、

$$\alpha = \{ \sum \sum (I1 - I2) \} / N \quad (\text{式 } 23)$$

のとき、E1は最小となる。この α は、ブロックマッチ

$$E1 = \sum \sum (I1 - I2)^2 - \{ \sum \sum (I1 - I2) \}^2 / N \quad (\text{式 } 24)$$

となるため、結局バイアス・ブロックマッチングでは式24を計算すればよい。あとは実施形態1同様の処理を経て最良マッチングを探索すればよい。

【0165】ここでは当然ながら、HVC濃度などRGB濃度以外の色空間による濃度を採用してもよい。2乗誤差の代わりに一般の1乗誤差、すなわち色の残差に基づいてブロックマッチングを行ってもよい。

【0166】(4)実施形態1のS13(対応点の初期位置の決定)では特徴点として、異時刻フレームt、t'、...において位置が安定的に変化している点を選んだが、ここでは選定基準を加重する。

【0167】図24は実施形態3で導入される特徴点の選定基準を示す図である。同図において、F10～F12の3枚がカメラ1によって撮影された異時刻フレーム、F20～F22の3枚がカメラ2によって撮影された異時刻フレームである。それぞれ左右2枚の組が同時刻フレームを示す。ここではある点Pに注目し、点Pの位置の異時刻フレーム間の動きをベクトルAnで、点Pの同時刻フレーム間の変位をベクトルBn(n:自然数)でそれぞれ表している。

【0168】以上の設定の下、実施形態3では、次の基準を満足する点を特徴点として選定する。

【0169】(a)ベクトルBnがほぼ一定であるか、ほぼ一定変化する

あるいは、この他、

(b)ベクトルAnがほぼ一定であるか、ほぼ一定変化する

の基準も追加し、(a)、(b)をとともに満足する点を特徴点として選定してもよい。

【0170】(b)は実施形態1で導入した条件に相当する。既述のごとく、多眼撮影では同時刻フレームのみから奥行き情報を求めることができる。しかしその前提となる映像間の対応関係の正確な把握は、これとは別問題であり、異時刻フレーム間の情報も積極的に利用すべ

像は、映像各部位の変位は別として、ほぼ同様の内容を含む。すなわちカメラの特性が近づくほど、式20のE1の値は小さくなる。この事実から逆に、 α はE1を最小にする値とすべきである。式20は、

【数14】

【数15】

ングの対象となる2つの領域間の各画素の色濃度差の平均値と言い換えられる。式23を式22に代入して計算すれば、

【数16】

きである。上記の2条件を同時に満たす点は、相当正確に追跡されていると考えられるため、二次元変位情報の抽出に対して重要な手掛かりを与える。ただし入力が静止映像の場合は、既知の動的計画法(ダイナミック・プログラミング)によって対応点を求めることもできる。

【0171】[工程2]奥行き情報の獲得

工程1で求められた映像各部位の変位から各部位の奥行き情報を導出する。多眼の場合、ある時刻tで図10の状態が実現されるので、以下、実施形態1の工程3の方法によって奥行き情報を獲得すればよい。

【0172】ここで注意すべきは、撮影カメラの位置関係が固定であるため、この関係およびカメラの倍率(焦点距離)が既知であれば、実施形態1で決まらなとされたスケールファクターcも含め、実施形態3では奥行き情報が正しく求まる。

【0173】[工程3]画像の生成

実施形態1の工程4(画像の生成)と同等の処理を行えばよい。

【0174】以上が実施形態3の概要である。実施形態3の場合、一般に非常に高い精度で奥行き情報を求めることができ、最終画像である視点変更映像等の精度も高くなる。

【0175】

【発明の効果】本発明によれば元の映像の奥行き情報から二次元表示用の画像を生成することができ、課題であった機能的・技術的制約が解消される。入力映像は単眼、多眼のいずれでもよく、適用範囲が広い。単眼の場合はシステム構成が簡素化でき、多眼の場合は奥行きの計算精度がさらに高くなる。

【0176】本発明によれば、複数の映像フレーム間における映像各部位の二次元的な位置の変位から奥行き情報を抽出することができる。

【0177】二次元的な位置の変位量に基づいて複数の映像フレームを選択する場合は、計算途上の誤差が減

る。

【0178】二次元的な位置の変位量を統計的に処理する場合は、選択の最適化、客観化が可能となる。

【0179】映像各部位の動きベクトルの分散が所定値よりも大きくなるように複数の映像フレームを選択する場合は、選択の客観化に加え、計算精度の保証が可能となる。分散が大きくなるような選択ができない場合、通常の映像表示等、他の表示で代替できる。

【0180】二次元的な位置の変化から映像各部位が現実の三次元空間において占める相対位置関係を導出する場合は、奥行きを正しく求めることができる。

【0181】映像各部位の三次元的な動きから三角測量の原理によって映像各部位の三次元空間における位置座標を算出する場合は、奥行きの決定を容易かつ正確に行うことができる。

【0182】本発明では、代表点と対応点という概念を導入したため、動き情報の把握が容易になる。

【0183】代表点と対応点の位置関係から他の異時刻フレームの対応点の位置を予想する場合は、対応点探索の計算量が削減される。

【0184】代表点のうち幾何的に特徴のある映像領域に関連するものについては、対応点の関連する映像領域についてもその特徴が維持されるため、対応点の位置精度が高い。

【0185】幾何的に特徴のある映像領域として直線を含む領域を採用する場合は、対応点の位置精度が改善される。

【0186】複数の映像フレーム間で映像の類似度が高い点を対応点とする場合は、対応点を正しく求めることができる。

【0187】映像の類似性のみならず、対応点間の相対位置の妥当性も評価する場合は、両評価の結果から対応点の位置精度を改善することができる。

【0188】前記各評価の結果が統合され、繰り返し計算がなされる場合は、対応点の位置精度を徐々に改善していくことができる。

【0189】前記類似性の評価をブロックマッチングによって行う場合は、容易かつ妥当な評価が可能となる。

【0190】ブロックマッチングでは、色濃度の n 乗誤差が計算されるため、映像に応じて実験等により最適の条件を設定することができる。

【0191】バイアスド・ブロックマッチングを行う場合は、特に多眼撮影された入力映像に対して良好な二次元表示画像の生成が可能となる。

【0192】色偏向定数を n 乗誤差総和が最小となるよう決めらる場合は、評価結果の信頼性が高まる。

【0193】色偏向定数として映像領域間の各画素の色濃度差の平均値を採用する場合は、2乗誤差総和が最小となり、評価結果の信頼性が高まる。

【0194】特徴点の対応点を優先的に決める場合は、

対応点の位置精度が高まる。

【0195】特徴点である対応点を補間することで非特徴点である対応点を決める場合は、非特徴点の対応点の位置精度も必然的に高くなる。

【0196】複数の映像フレーム間でブロックマッチングを行って特徴点を決める場合は、特徴点選定の客観化、選定結果の妥当性向上が可能となる。

【0197】異時刻フレームで位置の変化が安定して発生する点を特徴点とした場合、かかる点は精度よく追跡されているため、その位置精度が高い。

【0198】同時刻フレーム間における位置の変位が、その近傍時刻に撮影された同時刻フレーム間においてもほぼ一定となるか点を特徴点とする場合、多眼映像から二次元表示画像を良好に生成することができる。

【0199】奥行き情報に従って視点変更映像を生成する場合は、カメラの実際の移動が不要である。生成される画像も自然である。2眼映像から多視点映像を生成したり、滑らかなスローモーション映像を作ることとも可能となる。

【0200】前記奥行き情報を基準とし、映像の一部拡大・縮小を行う場合は、画像の自然さを失うことなく、強調表示等が可能となる。奥行きの小さい領域を拡大したり、逆に大きな領域を縮小する場合は特に効果的である。

【0201】これらの画像処理によって発生する画像端部の段差を修正する場合は、もとの映像同様、自然な画像枠を再生することができる。

【0202】奥行き情報に従って映像領域を切り出しを行う場合は、従来と全く異なる観点から相当正確な切り出しが可能となる。映像各部位のうち所定範囲の奥行きを持つ部位を切り出す場合は、映像の貼り付け等、画像加工処理が可能となる。

【図面の簡単な説明】

【図1】 実施形態1によって二次元表示画像を生成するための主な工程を示す図である。

【図2】 映像フレーム間の対応関係を検出するためのフローチャートである。

【図3】 基準フレーム t に代表点を設定する様子を示す図である。

【図4】 ブロックマッチングの様子を示す図である。

【図5】 仮の対応点 $P_t'(i, j)$ ごとに E の値を縦方向に示した模式図である。

【図6】 S12ステップの結果求められた対応点と代表点の関係を示す図である。

【図7】 対応点の相対位置を評価する原理を説明する図である。

【図8】 図6の対応点候補に対して本ステップの改善処理を行った結果を示す図である。

【図9】 ある点 P の画面上の移動と三次元空間での移動の対応を示す図である。

【図10】 カメラの三次元移動とある点Pの画面上の移動から点Pの三次元座標を導く原理を説明する図である。

【図11】 フレームtにおいて各代表点に数値が与えられた状態を示す図である。

【図12】 元の映像と視点変更映像の対応関係を示す図である。

【図13】 元の映像と視点変更映像の対応関係を示す図である。

【図14】 画像の一部を拡大して生成された画像を示す図である。

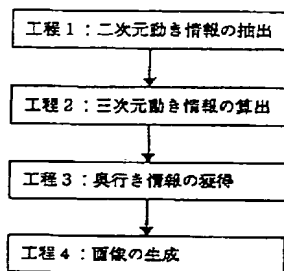
【図15】 図12の画像のうち、「家」を切り出して生成された画像を示す図である。

【図16】 実施形態1を実施するためのハードウェア構成例を示す図である。

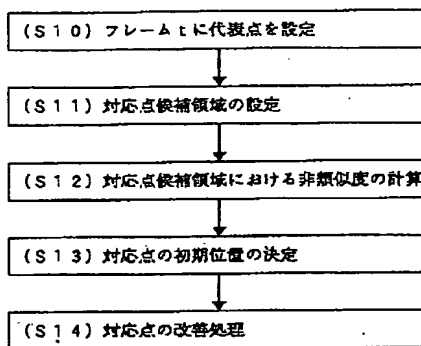
【図17】 フレームtの映像をディスプレイ上に表示した中間調画像の写真である。

【図18】 フレームt'の映像をディスプレイ上に表示した中間調画像の写真である。

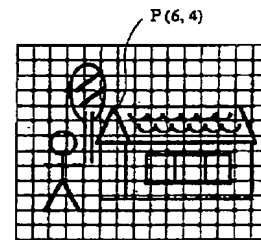
【図1】



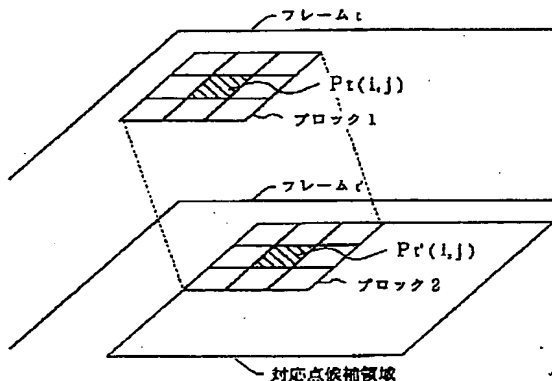
【図2】



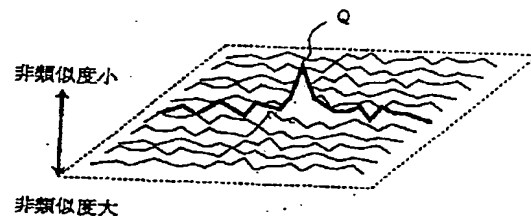
【図3】



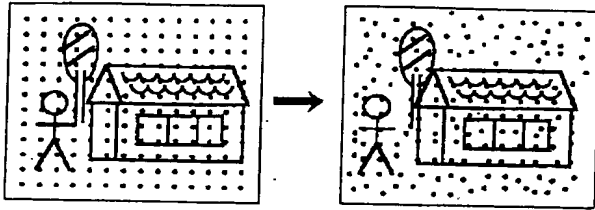
【図4】



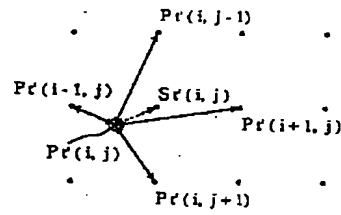
【図5】



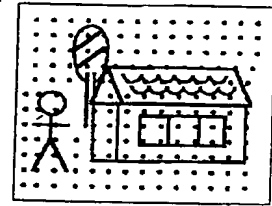
【図6】



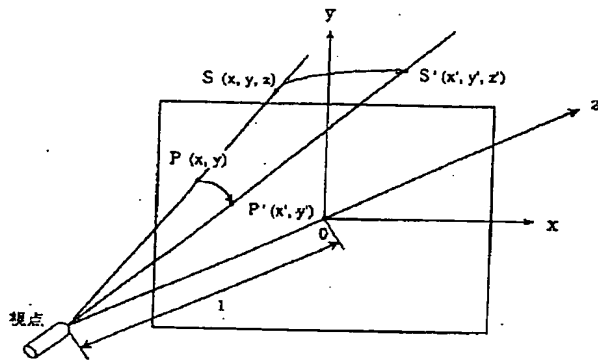
【図7】



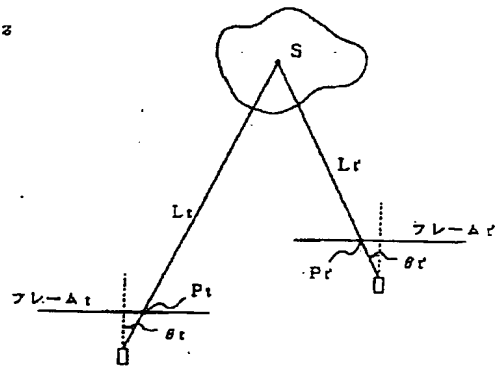
【図8】



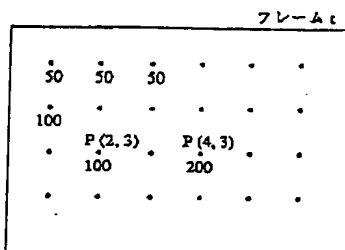
【図9】



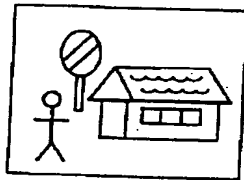
【図10】



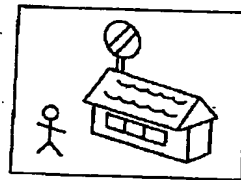
【図11】



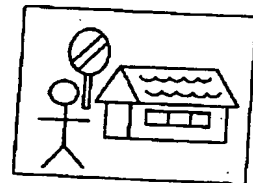
【図12】



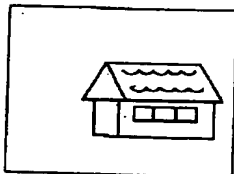
【図13】



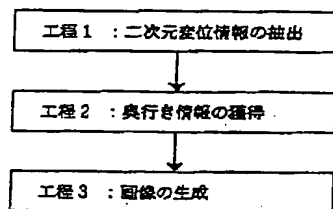
【図14】



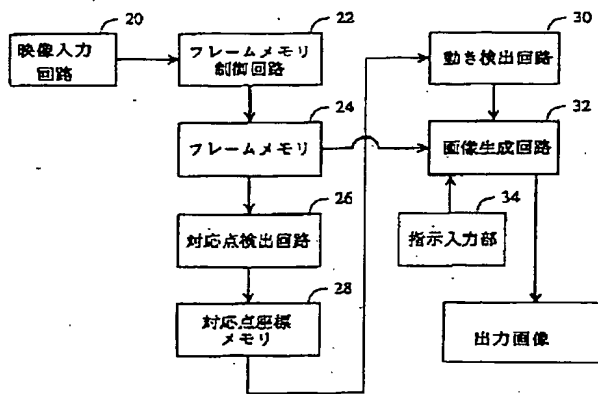
【図15】



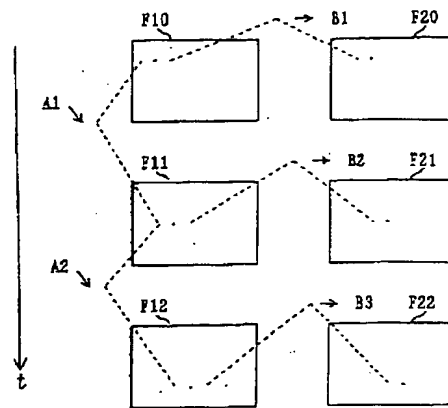
【図23】



【図16】



【図24】



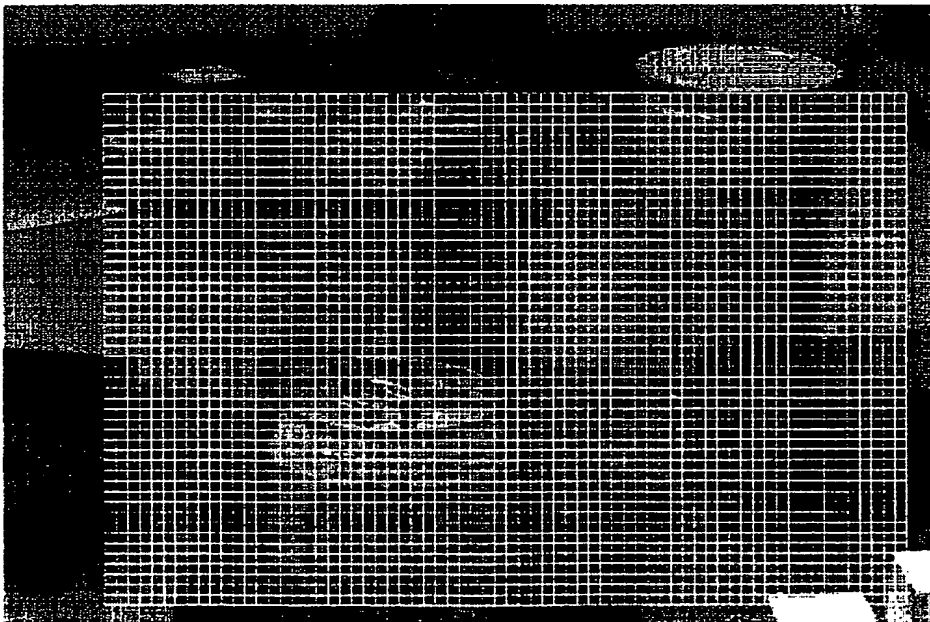
【図17】



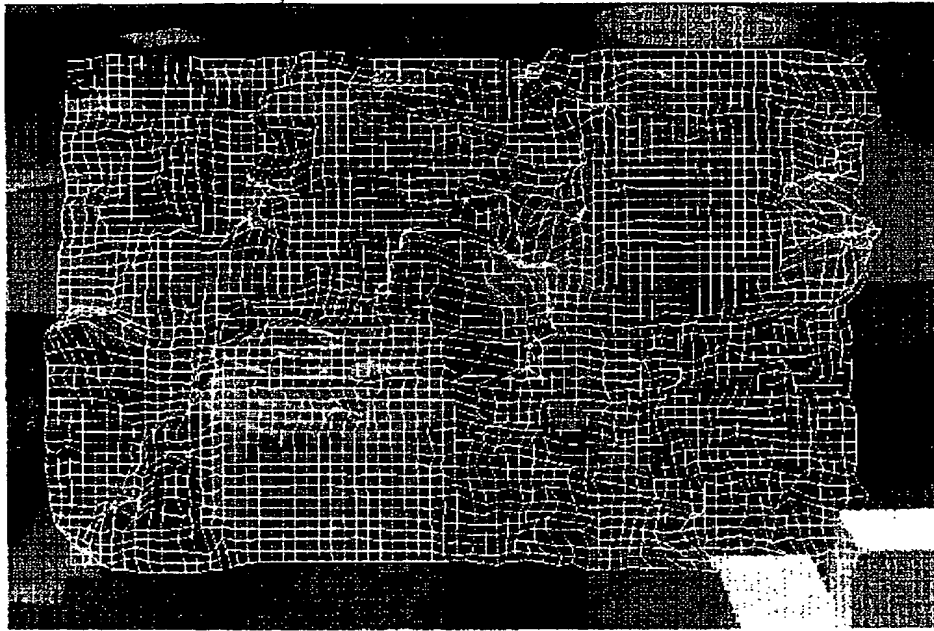
【図18】



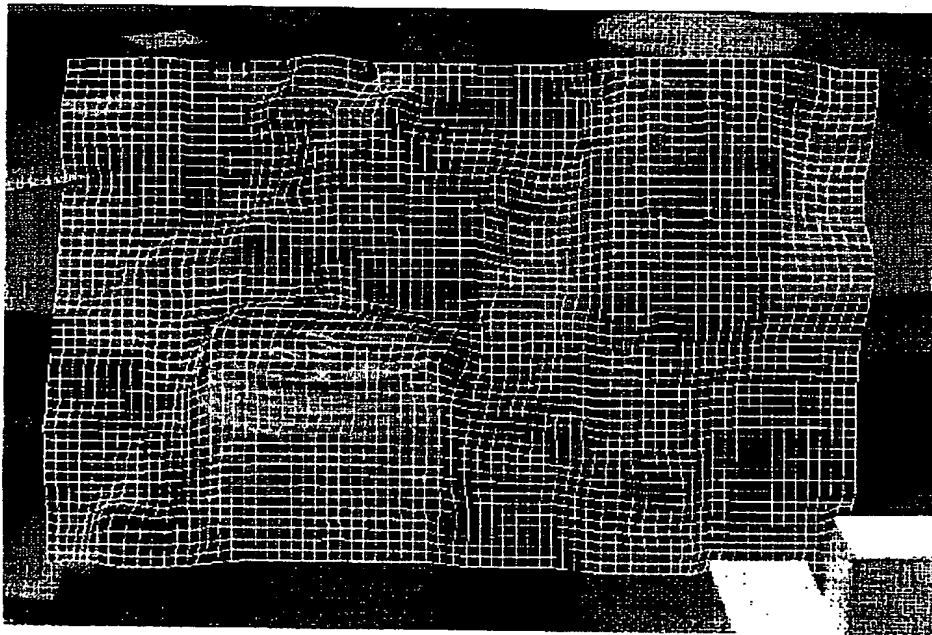
【図19】



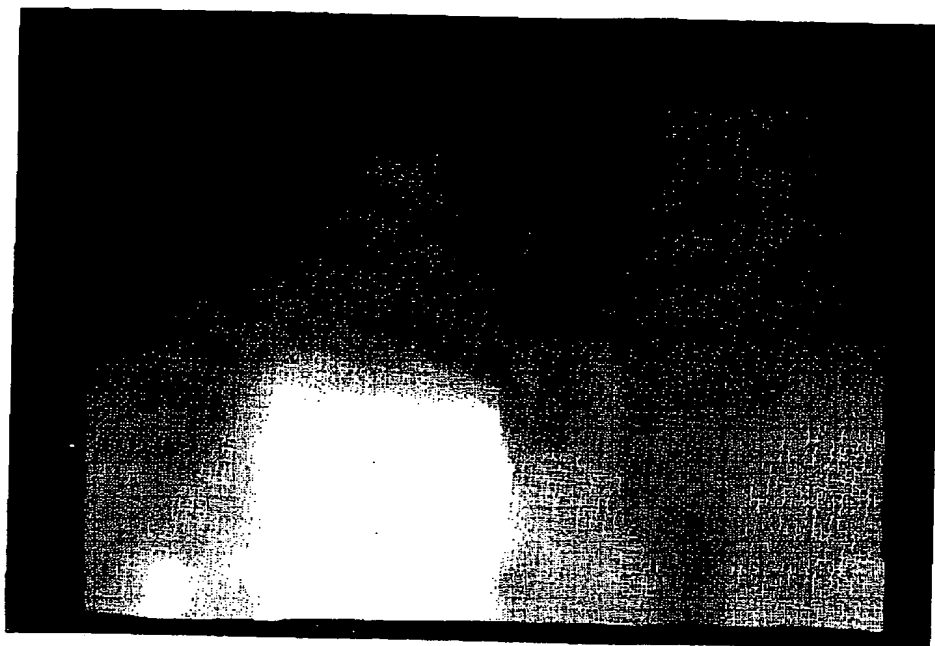
【図20】



【図21】



【図22】



フロントページの続き

(72)発明者 荒川 勉
大阪府守口市京阪本通2丁目5番5号 三
洋電機株式会社内